
Presentation of the Capstone IBM Coursera Project for Data Science

The battle of neighborhoods or the battle of
"*organic Food & Beverages*" at the close vicinity of
Berlin metro stations

Dr. B. Bayer, June 2021



Background

- ▶ Inhabitants of Germanys capital Berlin using mostly public transport system
- ▶ Even for daily commuting
- ▶ Increasing demand on healthy food in general and to be bought in Metro stations on the way to work

Question to be answered:

Which of Berlin metro stations are strategic for opening an "organic food and beverage" business?

Data used

Wikipedia

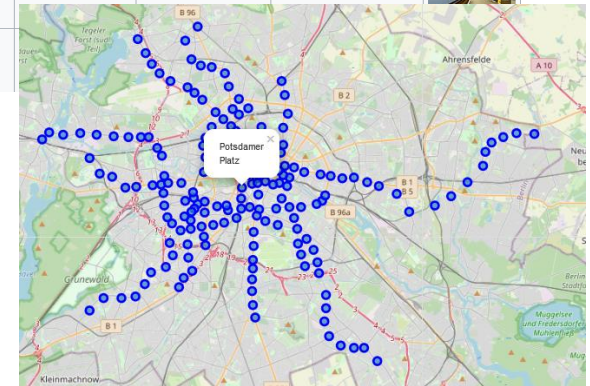
for coordinates of Metro stations

https://de.wikipedia.org/wiki/Liste_der_Berliner_U-Bahnhöfe

Stationen [Bearbeiten] [Quelltext bearbeiten]

Inhaltsverzeichnis A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

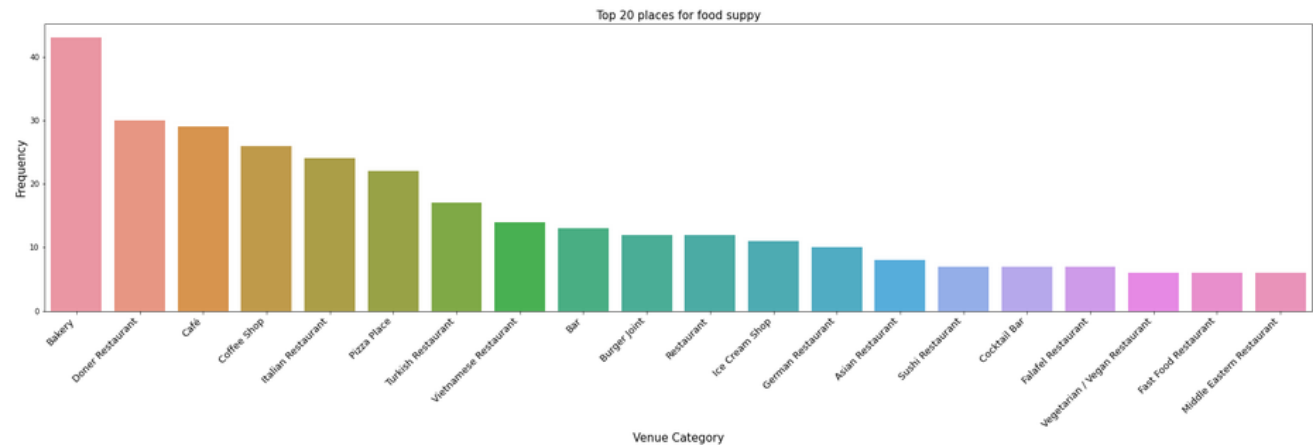
Bahnhof (Kurz- Karte)	Linie	Eröffnung	Lage	Ortsteil	Umsstieg	Denkmal	Anmerkungen	Sehenswürdigkeiten	Bild
Adenauerplatz (Ado) 52° 29' 59" N, 13° 18' 26" O	U7	28. Apr. 1978	Tunnel	Charlottenburg		–			
Afrikanische Straße (Afr) 52° 33' 38" N, 13° 20' 3" O	U6	3. Mai 1956	Tunnel	Wedding		–			
Alexanderplatz (A) 52° 31' 17" N, 13° 24' 48" O	U2	1. Juli 1913	Tunnel	Mitte					



Foursquare venue data

- Search radius of 100 m
- 842 venue entries. 184 relevant for food industry

Top 20 venues: Bakery and Doner



Methodology

- ▶ Feature extracting
 - Input Foursquare venue data

- ▶ Clustering
 - The stations were clustered based on a set of similar characteristics or features, i.e., their surrounding venues.
 - *K*-Means clustering: an unsupervised machine learning algorithm that creates homogeneous subgroups/clusters from unlabeled data such that data points in each cluster are as similar as possible to each other according to a similarity measure (e.g., Euclidian distance).

Results and recommendation

► Color coded map

Cluster 0: wine red

Cluster 1: dark blue

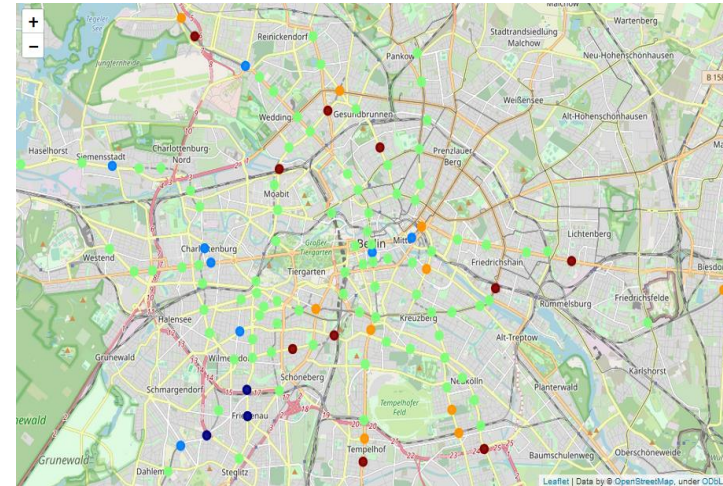
Cluster 2: brighter blue

Cluster 3: bright green

Cluster 4: orange

► Overview cluster:

Cluster Label	Member	Common Venue
0	13	Bakeries
1	3	Mexican and Wine Shops
2	13	Italian and Wine Shops
3	92	Coffee/Cafe, Pizza, Turkish food and Wine Shops
4	14	Doner restaurants and Wine Shops



► Top most common venues for cluster 3 with 92 members:

Coffee/Cafe, Pizza, Turkish food and Wine Shops

► Note: „Healthy“ food categories (such as vegetarian/vegan) are not among the top ratings.

→ **High potential** to start opening organic food and beverage places on each Metro/tram station of Berlin